



PATENT ABSTRACTS OF JAPAN

(11) Publication number: 2000187495 A

(43) Date of publication of application: 04.07.2000

(51) Int. Cl. G10L 13/08
G10L 13/06(21) Application number: 10363283
(22) Date of filing: 21.12.1998(71) Applicant: NEC CORP
(72) Inventor: INAGAKI KEIKO
MITOME YUKIO(54) METHOD AND DEVICE FOR SYNTHESIZING
SPEECH, AND RECORDING MEDIUM WHERE
SPEECH SYNTHESIS PROGRAM IS
RECORDED

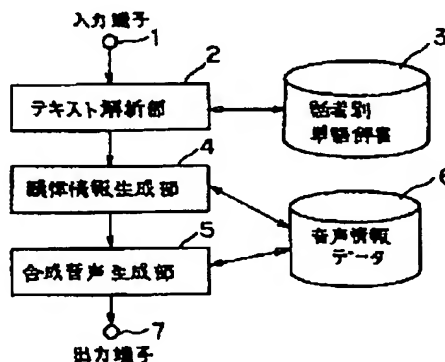
(57) Abstract:

PROBLEM TO BE SOLVED: To generate utterance information considering individuality of a speaker.

SOLUTION: An entry, a part of speech, reading, and an accent position of each word are stored in a word dictionary 3 by speaker. A rhythm pattern and a speech waveform extracted from the sentences and words which a speaker, who made the word dictionary 3 by speaker, uttered are stored in a speech information data 6 with information on text, pronunciation, and label. A text analysis part 2 generates pronunciation information consisting of reading, an accent, and a pose by using the word dictionary 3 by speaker to an input text inputted from an input terminal 1. A rhythm information generating part 4 generates patterns of the intonation and the rhythm from the rule and pattern stored in a speech information data 6 based on the

pronunciation information generated in the text analysis part 2. A synthetic speech generating part 5 generates a synthetic speech by using the speech pattern stored in the speech information data 6 based on the pronunciation information and rhythm information, and outputs the synthesized speech from an output terminal 7.

COPYRIGHT: (C)2000,JPO



BEST AVAILABLE COPY

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2000-187495

(P2000-187495A)

(43) 公開日 平成12年7月4日(2000.7.4)

(51) Int. Cl.	識別記号	F I	キーワード(参考)
G 1 0 L 13/08		G 1 0 L 3/00	H 5 D 0 4 5
13/06		5/04	F

審査請求 有 請求項の範囲 OL (全 9 頁)

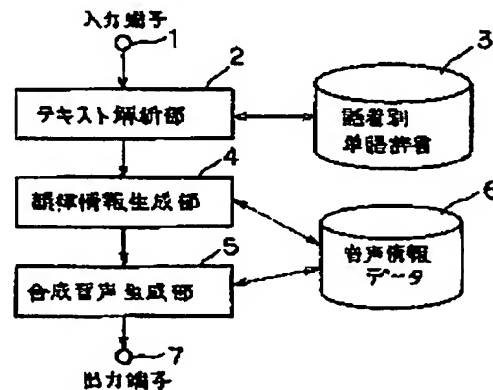
(21) 出願番号	特願平10-963283	(71) 出願人	000004237 日本電気株式会社 東京都港区芝五丁目7番1号
(22) 出願日	平成10年12月21日(1998.12.21)	(72) 発明者	藤田 敦子 東京都港区芝五丁目7番1号 日本電気株式会社内
		(72) 発明者	三留 幸夫 東京都港区芝五丁目7番1号 日本電気株式会社内
		(74) 代理人	100100888 弁護士 渡辺 勝 (外3名) Pターム(参考) BD045 AA07 AA09

(54) 【発明の名称】 音声合成方法、装置、および音声合成プログラムを記録した記録媒体

(57) 【要約】

【課題】 話者の個性を考慮した発音情報を生成する。

【解決手段】 話者別単語辞書3には、各単語の見出し、品詞、読みとアクセント位置が記憶されている。音声情報データ6には、話者別単語辞書3を作成した話者が発声した文や単語から抽出された韻律パターンと音声波形が、テキスト情報と発音情報、ラベル情報とともに記憶されている。テキスト解析部2は入力端子1から入力された入力テキストに対して話者別単語辞書3を用いて読みとアクセント、ポーズからなる発音情報を生成する。韻律情報生成部4はテキスト解析部2で生成された発音情報に基づいて、イントネーションやリズムのパターンを、音声情報データ6に記憶されている規則やパターンから生成する。合成音声生成部5は、発音情報と韻律情報に基づいて音声情報データ6に記憶されている音声パターンを用いて合成音声を生成し、出力端子7から出力する。



(2)

特開2000-187495

1

【特許請求の範囲】

【請求項1】 入力テキストに対し、話者の話すとおりの読みとアクセントを持つ話者別単語辞書を用いて音韻情報を生成するテキスト解析段階と、
前記入力テキストの韻律情報を生成する韻律情報生成段階と、

前記音韻情報と前記韻律情報に基づいて合成音声を生成する合成音声生成段階とを有する音声合成方法。

【請求項2】 前記話者別単語辞書が話者毎に存在し、前記テキスト解析段階で、入力された話者情報から前記話者別単語辞書を切り替える、請求項1記載の方法。

【請求項3】 入力テキストに対し、国語辞典に記述されている通常の読みを持つ単語辞書と、その読みを話者特有の読みに変換する読み変換手段を用いて音韻情報を生成するテキスト解析段階と、

前記入力テキストの韻律情報を生成する韻律情報生成段階と、

前記音韻情報と前記韻律情報に基づいて合成音声を生成する合成音声生成段階とを有する音声合成方法。

【請求項4】 入力テキストに対し、国語辞典に記述されている通常の読みを持つ単語辞書と、その読みを話者特有の読みに変換する、話者毎の読み変換手段を用いて、話者毎に前記読み変換手段を切り替えて音韻情報を生成するテキスト解析段階と、

前記入力テキストの韻律情報を生成する韻律情報生成段階と、

前記音韻情報と前記韻律情報に基づいて合成音声を生成する合成音声生成段階とを有する音声合成方法。

【請求項5】 入力テキストに対して、話者の話すとおりの読みとアクセントを持つ話者別単語辞書と、該話者別単語辞書の読みから話者特有の読みに変換する読み変換手段を用いて、音韻情報を生成するテキスト解析段階と、

前記入力テキストの韻律情報を生成する韻律情報生成段階と、

前記音韻情報と前記韻律情報に基づいて合成音声を生成する合成音声生成段階とを有する音声合成方法。

【請求項6】 入力テキストに対して、話者の話すとおりの読みとアクセントを持つ、話者毎の話者別単語辞書と、該話者別単語辞書の読みから話者特有の読みに変換する、話者毎の読み変換手段を用いて、話者毎に話者別単語辞書と読み変換手段を切り替えて、音韻情報を生成するテキスト解析段階と、

前記入力テキストの韻律情報を生成する韻律情報生成段階と、

前記音韻情報と前記韻律情報に基づいて合成音声を生成する合成音声生成段階とを有する音声合成方法。

【請求項7】 話者の話すとおりの読みとアクセントを持つ話者別単語辞書と、

入力テキストに対して前記話者別単語辞書を用いて音韻

2

情報を生成するテキスト解析手段と、

前記入力テキストの韻律情報を生成する韻律情報生成手段と、

前記音韻情報と前記韻律情報に基づいて合成音声を生成する合成音声生成手段を有する音声合成装置。

【請求項8】 話者の話すとおりの読みとアクセントを持つ、話者毎の話者別単語辞書と、
入力された話者情報から前記話者別単語辞書を切り替え、入力テキストに対して該話者別単語辞書を用いて音韻情報を生成するテキスト解析手段と、

前記入力テキストの韻律情報を生成する韻律情報生成手段と、

前記音韻情報と前記韻律情報に基づいて合成音声を生成する合成音声生成手段を有する音声合成装置。

【請求項9】 国語辞典に記述されている通常の読みを持つ単語辞書と、

前記単語辞書の読みを話者特有の読みに変換する読み変換手段と、

入力テキストに対して前記単語辞書と読み変換手段を用いて音韻情報を生成するテキスト解析手段と、

前記入力テキストの韻律情報を生成する韻律情報生成手段と、

前記音韻情報と前記韻律情報に基づいて合成音声を生成する合成音声生成手段を有する音声合成装置。

【請求項10】 国語辞典に記述されている通常の読みを持つ単語辞書と、

前記単語辞書の読みを話者特有の読みに変換する、話者毎の読み変換手段と、

入力された話者情報から前記読み変換手段を切り替え、入力テキストに対して前記単語辞書と該読み変換手段を用いて音韻情報を生成するテキスト解析手段と、

前記入力テキストの韻律情報を生成する韻律情報生成手段と、

前記音韻情報と前記韻律情報に基づいて合成音声を生成する合成音声生成手段を有する音声合成装置。

【請求項11】 話者の話すとおりの読みとアクセントを持つ、話者別単語辞書と、

該話者別単語辞書の読みから話者特有の読みに変換する読み変換手段と、

入力テキストに対して、前記話者別単語辞書と前記読み変換手段を用いて、音韻情報を生成するテキスト解析手段と、

前記入力テキストの韻律情報を生成する韻律情報生成手段と、

前記音韻情報と前記韻律情報に基づいて合成音声を生成する合成音声生成手段とを有する音声合成装置。

【請求項12】 話者の話すとおりの読みとアクセントを持つ、話者毎の話者別単語辞書と、

該話者別単語辞書の読みから話者特有の読みに変換する、話者毎の読み変換手段と、

入力テキストに対して前記話者別単語辞書を用いて音韻

50

(3)

特開2000-187495

3

入力された話者情報から前記話者別単語辞書と前記読み変換手段を切り替え、入力テキストに対して、前記話者別単語辞書と前記読み変換手段を用いて、音韻情報を生成するテキスト解析手段と、

前記入力テキストの韻律情報を生成する韻律情報生成手段と、

前記音韻情報と前記韻律情報に基づいて合成音声を生成する合成音声生成手段とを有する音声合成装置、

【請求項13】 入力テキストに対し、話者の話すとおりの読みとアクセントを持つ話者別単語辞書を用いて音韻情報を生成するテキスト解析手段と、

前記入力テキストの韻律情報を生成する韻律情報生成手段と、

前記音韻情報と前記韻律情報に基づいて合成音声を生成する合成音声生成手段をコンピュータに実行させるための音声合成プログラムを記録した記録媒体、

【請求項14】 前記話者別単語辞書が話者毎に存在し、前記テキスト解析手段で、入力された話者情報から前記話者別単語辞書を切り替える、請求項13記載の記録媒体、

【請求項15】 入力テキストに対し、国語辞典に記述されている通常の読みを持つ単語辞書と、その読みを話者特有の読みに変換する読み変換手段を用いて音韻情報を生成するテキスト解析手段と、

前記入力テキストの韻律情報を生成する韻律情報生成手段と、

前記韻律情報と前記韻律情報に基づいて合成音声を生成する合成音声生成手段とをコンピュータに実行させるための音声合成プログラムを記録した記録媒体、

【請求項16】 入力テキストに対して、国語辞典に記述されている通常の読みを持つ単語辞書と、その読みを話者特有の読みに変換する、話者毎の読み変換手段を用いて、話者毎に前記読み変換手段を切り替えて音韻情報を生成するテキスト解析手段と、

前記入力テキストの韻律情報を生成する韻律情報生成手段と、

前記音韻情報と前記韻律情報に基づいて合成音声を生成する合成音声生成手段とをコンピュータに実行させるための音声合成プログラムを記録した記録媒体、

【請求項17】 入力テキストに対して、話者の話すとおりの読みとアクセントを持つ話者別単語辞書と、該話者別単語辞書の読みから話者特有の読みに変換する読み変換手段を用いて、音韻情報を生成するテキスト解析手段と、

前記入力テキストの韻律情報を生成する韻律情報生成手段と、

前記音韻情報と前記韻律情報に基づいて合成音声を生成する合成音声生成手段とをコンピュータに実行させるための音声合成プログラムを記録した記録媒体、

【請求項18】 入力テキストに対し、話者の話すとお

4

りの読みとアクセントを持つ、話者毎の話者別単語辞書と、該話者別単語辞書の読みから話者特有の読みに変換する、話者毎の読み変換手段を用いて、話者毎に話者別単語辞書と読み変換手段を切り替えて、音韻情報を生成するテキスト解析手段と、

前記入力テキストの韻律情報を生成する韻律情報生成手段と、

前記音韻情報と前記韻律情報に基づいて合成音声を生成する合成音声生成手段とをコンピュータに実行させるための音声合成プログラムを記録した記録媒体、

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、テキストに基づいて合成音声を生成する音声合成方法および装置に関する。

【0002】

【従来の技術】従来、日本語テキストを音声に変換する音声合成装置として様々な装置が開発されている。現在では、より高品質な合成のために、合成音声の個人性の向上が多く望まれている。個人性とは、合成音声が固一的でなく、その人らしさを再現するものであり、あたかも本人が話しているかのような合成音声を生成するためには重要な要素である。この合成音声の個人性を向上させる合成装置の従来例として特開平8-87297（従来例1）がある。

【0003】図7は、この従来例1の音声合成装置の概略構成を示すブロック図であり、任意の入力テキストを合成音声に出力する場合を示している。この従来の音声合成装置は音声情報検索部21と音声情報データベース22と合成音声生成部23と合成音声生成規則24とから構成されている。

【0004】音声情報検索部21は、テキストまたは発音記号列が入力されると、自然音声を分析し、抽出した音声特徴量およびこれに対応する発音内容を格納した音声情報データベース22中に、入力テキストまたは発音記号列に一致する発音内容が存在するか否かを検索し、一致する発音内容が存在する場合は、これを合成音声生成部23へ渡し、音声情報に応じた処理を施して合成音声を生ずる。また、音声情報検索部21は一致する発音内容が存在しない場合には、入力テキストまたは入力発音記号列をそのまま合成音声生成部23へ渡し、合成音声生成部23は合成音声生成規則24に基づいて合成音声を生ずる。このように、音声情報データベース22にあらかじめ自然音声記憶させておき、発音させたい内容と同じものが音声情報データベース22にある場合には、音声合成を行わず自然音声を出力することで、元の話者の個人性を含む合成音声を生じている。

【0005】また、他の従来例として、特開平10-153998（従来例2）がある。これは、自然音声を補助情報として利用し、規則合成により音声を作成する装

(4)

特開2000-187495

5

6

置で、単語辞書を参照して入力テキストの解析により得た単語系列の各単語の音素系列に対する韻律情報を設定し、音素波形辞書を参照して各単語の音素系列から音素波形系列を求めている。また、入力音素から韻律情報を抽出し、設定された韻律情報と抽出された韻律情報のいずれかを選択し、その選択された韻律情報により音素波形系列を創制し、合成音声を生成している。このように自然音声から抽出された韻律情報には、元の話者のイントネーションや発声速度などの個人性が含まれている。

【0006】

【課題を解決しようとする課題】従来例1の音声合成装置では、自然音声から抽出した音素波形をそのまま使用することで、合成音声の個人性を向上させている。また、従来例2の音声合成装置では、自然音声から抽出した韻律パターンをそのまま用いて、元の話者の話し方を再現することで個人性を向上させている。しかし、従来例1、2では、一致する文が記憶したデータにない場合には、規則によりそれらを生成するため、自然音声から抽出したパターンを使った合成音声と比較すると個人性は劣化してしまう。従来例1、2において個人性を向上させるためには、膨大なデータが必要であり、すべての文を網羅するパターンを用意することは困難である。また、従来例1、2では、入力テキストを解析し、その発音情報を単語辞書を用いて生成するが、この単語辞書は、どの話者データに対しても共通であるため、テキストの解析結果も同じである。しかし、たとえ同一の文であっても、実際には発音情報が話者により微妙に異なることは多々ある。この違いは、従来例1や2の方法では、対処することができない。合成音声をより一層元の話者の発声に近づけるためには、この発音情報を生成する時にも個人性を考慮する必要がある。

【0007】本発明の目的は、合成音声の個人性を向上させるために、話者の個人性を考慮した発音情報を生成することができる音声合成方法および装置を提供することにある。

【0008】

【課題を解決するための手段】本発明の第1の音声合成方法は、入力テキストに対し、話者の話すおりの読みとアクセントを持つ話者別単語辞書を用いて音韻情報を生成するテキスト解析段階と、入力テキストの韻律情報を生成する韻律情報生成段階と、音韻情報と前記韻律情報に基づいて合成音声を生成する合成音声生成段階とを有する。

【0009】本発明の第1の音声合成装置は、話者の話すおりの読みとアクセントを持つ話者別単語辞書と、入力テキストに対して前記話者別単語辞書を用いて音韻情報を生成するテキスト解析手段と、入力テキストの韻律情報を生成する韻律情報生成手段と、音韻情報と韻律情報に基づいて合成音声を生成する合成音声生成手段を

有する。

【0010】本発明の第2の音声合成方法は、第1の音声合成方法において、話者別単語辞書が話者毎に存在し、テキスト解析段階で、入力された話者情報から話者別単語辞書を切り替える。

【0011】本発明の第2の音声合成装置は、話者の話すおりの読みとアクセントを持つ、話者毎の話者別単語辞書と、入力された話者情報から話者別単語辞書を切り替え、入力テキストに対して該話者別単語辞書を用いて音韻情報を生成するテキスト解析手段と、入力テキストの韻律情報を生成する韻律情報生成手段と、音韻情報と韻律情報に基づいて合成音声を生成する合成音声生成手段を有する。

【0012】本発明の第3の音声合成方法は、入力テキストに対し、国語辞典に記述されている通常の読みを持つ単語辞書と、その読みを話者特有の読みに変換する読み変換手段を用いて音韻情報を生成するテキスト解析段階と、入力テキストの韻律情報を生成する韻律情報生成段階と、音韻情報と韻律情報に基づいて合成音声を生成する合成音声生成段階とを有する。

【0013】本発明の第3の音声合成装置は、国語辞典に記述されている通常の読みを持つ単語辞書と、単語辞書の読みを話者特有の読みに変換する読み変換手段と、入力テキストに対して単語辞書と読み変換手段を用いて音韻情報を生成するテキスト解析手段と、入力テキストの韻律情報を生成する韻律情報生成手段と、音韻情報と韻律情報に基づいて合成音声を生成する合成音声生成手段とを有する。

【0014】本発明の第4の音声合成方法は、入力テキストに対して、国語辞典に記述されている通常の読みを持つ単語辞書と、その読みを話者特有の読みに変換する、話者毎の読み変換手段を用いて、話者毎に読み変換手段を切り替えて音韻情報情報を生成するテキスト解析段階と、入力テキストの韻律情報を生成する韻律情報生成段階と、音韻情報と韻律情報に基づいて合成音声を生成する合成音声生成段階とを有する。

【0015】本発明の第4の音声合成装置は、国語辞典に記述されている通常の読みを持つ単語辞書と、単語辞書の読みを話者特有の読みに変換する、話者毎の読み変換手段と、入力された話者情報から読み変換手段を切り替え、入力テキストに対して単語辞書と読み変換手段を用いて音韻情報を生成するテキスト解析手段と、入力テキストの韻律情報を生成する韻律情報生成手段と、音韻情報と韻律情報に基づいて合成音声を生成する合成音声生成手段を有する。

【0016】本発明の第5の音声合成方法は、入力テキストに対して、話者の話すおりの読みとアクセントを持つ話者別単語辞書と、該話者別単語辞書の読みから話者特有の読みに変換する読み変換手段を用いて、音韻情報を生成するテキスト解析段階と、入力テキストの韻律

(5)

特開2000-187495

7

8

情報を生成する韻律情報生成段階と、音韻情報と韻律情報に基づいて合成音声生成する合成音声生成段階とを有する。

【0017】本発明の第5の音声合成装置は、話者の話すおりの読みとアクセントを持つ話者別単語辞書と、該話者別単語辞書の読みから話者特有の読みに変換する読み変換手段と、入力テキストに対して、話者別単語辞書と読み変換手段を用いて、音韻情報を生成するテキスト解析手段と、入力テキストの韻律情報を生成する韻律情報生成手段と、音韻情報と韻律情報に基づいて合成音

声を生成する合成音声生成手段とを有する。
【0018】本発明の第6の音声合成方法は、入力テキストに対して、話者の話すおりの読みとアクセントを持つ、話者毎の話者別単語辞書と、該話者別単語辞書の読みから話者特有の読みに変換する、話者毎の読み変換手段を用いて、話者毎に話者別単語辞書と読み変換手段を切り替えて、音韻情報を生成するテキスト解析段階と、入力テキストの韻律情報を生成する韻律情報生成段階と、音韻情報と韻律情報に基づいて合成音声を生成する合成音声生成段階とを有する。

【0019】本発明の第6の音声合成装置は、話者の話すおりの読みとアクセントを持つ、話者毎の話者別単語辞書と、該話者別単語辞書の読みから前記話者別単語辞書と話者特有の読みに変換する、話者毎の読み変換手段と、入力された話者情報から読み変換手段を切り替え、入力テキストに対して、話者別単語辞書と読み変換手段を用いて、音韻情報を生成するテキスト解析手段と、入力テキストの韻律情報を生成する韻律情報生成手段と、音韻情報と韻律情報に基づいて合成音声を生成する合成音声生成手段とを有する。

【0020】話者の発音に特化した韻律や変換手段を持つことで、合成音声の発音情報が元の話者と同じものになる。

【0021】

【発明の実施の形態】次に、本発明の実施の形態について図面を参照して詳細に説明する。図1を参照すると、本発明の第1の実施の形態は、テキストを入力する入力端子1と、入力テキストと話者別単語辞書3を用いて読みとアクセント、ポーズからなる発音情報を生成するテキスト解析部2と、テキスト解析部2で生成された発音情報に基づいて、イントネーションやリズムのボタンを、音声情報データ6に記憶されている規則やボタンから生成する韻律情報生成部4と、発音情報と韻律情報に基づいて音声情報データ6に記憶されている音声パタ

ンを用いて合成音声を生成する合成音声生成部5と、合成音声生成部5で生成された合成音声を出力する出力端子7で構成されている。話者別単語辞書3には、各単語の見出し、品詞、読みとアクセント位置が記憶されている。この話者別単語辞書3を作成する方法としては、例えば以下のような方法がある。まず、従来の音声合成に使われるような標準的な日本語の発音の入った辞書を用意しておく。次に、合成に使う話者が本人の発音に近い発音辞書を作成するために、文や単語を発声したところから抽出した発音情報を単語毎に分割し、該当する辞書の発音情報と置き換えていく。または、音声合成に使用する話者が、音声情報データベースを作成するために収録した文や単語のデータがある場合には、そこから抽出した発音情報を発音辞書に反映させる。このようにして、話者特有の発音情報を持つ辞書を作成していく。この話者別単語辞書3には、合成に使う話者が実際に発声した音声から発音情報を抽出し、記憶しておくため、従来の発音辞書よりもより個性が反映されたものとなる。音声情報データ6には、話者別単語辞書3を作成した話者が発声した文や単語から抽出された韻律ボタンと音声波形が、テキスト情報と発音情報、ラベル情報とともに記憶されている。次に、本実施形態の動作について詳細に説明する。入力端子1からは、例えば「これを校正してください。」という文がキーボード等の入力手段を用いて入力される。テキスト解析部2では、上記入力文に対し話者別単語辞書3を用いて、(これお、こーぜーしてください%い/(「;」は、ポーズ位置、"%は、アクセント位置を示す。)のような読みとアクセント情報、ポーズ位置が入った発音情報が生成される。韻律情報生成部4では、テキスト解析部3で生成された発音情報にマッチする韻律ボタンを音声情報データ6から選択、編集し、韻律ボタンを生成し、発音情報と共に合成音声生成部5に供給する。合成音声生成部5では、発音情報と韻律ボタンを基に音声情報データ6より音声波形を選択、編集し、合成音声を作成し、出力端子7より出力する。表1、表2は話者別単語辞書3の一例であり、話者Aと話者Bの発声を基にそれぞれ作られた辞書であり、表3は従来の音声合成に使われた発音辞書で、アクセント辞典や国語辞典等の記述によるものである。表1、表2、表3の単語辞書には、各単語に対し、発音情報と品詞が記憶されている。

【0022】

【表1】

(6)

特選2000-187495

9

10

見出し	発音情報	品詞
英語	えいご	名詞
西部	せいぶ	名詞
服	ふく	名詞
靴	くつ	名詞

[0023]

※ ※【表2】

見出し	発音情報	品詞
英語	えーご	名詞
西部	せーぶ	名詞
服	ふく	名詞
靴	くつ	名詞

[0024]

※ ※【表3】

見出し	発音情報	品詞
英語	えーご	名詞
西部	せーぶ	名詞
服	ふく%	名詞
靴	くつ%	名詞

これら3つの単語辞書を比較すると、同じ単語の発音情報が辞書によって異なっていることがわかる。例えば、「英語」という単語は、表1の話者Aの辞書には、「えいご」と「い」を比較的にはっきり発音するように記述されているが、表2の話者Bや表3では「えーご」と「え」の長音化のように発音するように記述されている。また、「服」は、表1の話者Aの辞書では、「ふ

にアクセント核のある1型アクセントだが、表2の話者Bではアクセントがない0型であり、表3の辞書では、語尾につく2型となっている。このように、表1から表3までの辞書のうちいずれかを使うかにより、合成音声の発音情報は変化する。言い換えれば、ここに話者特有の発音情報を記述しておくことで、話者の個性を表現することが可能となる。次に、本発明の第2の実施の

(7)

特開2000-187495

11

形態について図面を参照して説明する。図2の本発明の第2の実施形態は、第1の実施形態にある話者別単語辞書3のかわりに、単語辞書8と読み変換部9を有する点で第1の実施形態と異なる。単語辞書8の構造は話者別単語辞書3と同じであるが、発音情報がある特定の話者の発音データを基に作成されたのではなく、国語辞典やアクセント辞典を基に作成されたものである点で話者別単語辞書3と異なっている。また、読み変換部9は、単語辞書8で選択された単語の発音情報を規則によって変換し、テキスト解析部2へ渡す。このように単語辞書8と読み変換部9を有することで、第2の実施形態でも、話者に特化した読みを供給することが可能となる。読み変換部9に記憶されている変換規則は、例えば、音声台成に使用する話者が、音声情報データベースを作成するために収録した文や単語のデータを用いて作成することができる。この規則には、ある特定の単語の変換規則を記述することもできるが、いくつか変例から一般化した規則を記述してもよい。例えば、「総」を/えー/、「蚊」を/かー/、「子」を/こー/と長音化する話者がいたとする。このような話者は、他の1モーラの単語「歯」や「名」等も同様に長音化させる傾向がある。そこで、音声台成に使用する話者の発音情報を基に、話者の発音情報がない単語も変換させるような規則を記述しておくこともできる。この場合、「総」、「蚊」、「子」は長音化させる」というのが、特定の単語の変換規則にあたり、「1モーラの単語は長音化する」という規則が一般化された変換規則にあたる。このように、変換規則を一般化しておくことで、音声台成に使用する話者の発音情報がない単語でも話者特有の発音情報を変換することができる。

【0025】次に、本発明の第3の実施形態について図面を参照して説明する。図3の本発明の第3の実施形態は、第1の実施形態にある話者別単語辞書3と、第2の実施形態にある読み変換部9の両方を有する点の特徴とする。入力部1より入力されたテキストはテキスト解析部2において、まず、話者別単語辞書3を用いて単語の発音情報が選択される。選択された発音情報は、読み変換部9により変換規則にマッチする単語がある場合には、発音情報が変換され、テキスト解析部2へ発音情報として戻される。話者別単語辞書3は、話者が文や単語を発した音声から抽出した発音情報を基に作成されるため、話者が発生していない単語の発音情報は標準的な日本語のままである。一方、読み変換部9には、話者が発生した音声から抽出した発音情報を基に、それらが一般化された変換規則が格納されている。この一般化された変換規則を使用することで、話者が発生していないため話者別単語辞書3では変換できない発音情報を話者性を考慮した発音情報に変換することができ

る。

【0026】次に、本発明の第4の実施形態について

12

図面を参照して説明する。図4の本発明の実施形態は、第1の実施形態にある話者別単語辞書3と音声情報データ6の組を話者A,B,C毎に話者別単語辞書3Aと音声情報データ6A、話者別単語辞書3Bと音声情報データ6B、話者別単語辞書3Cと音声情報データ6Cを備えたものである。話者情報に台わせて、対応する話者別単語辞書と音声情報データを選択する。話者別単語辞書と音声情報データには、あらかじめ話者情報を格納しておく。このように話者情報に台わせて、話者別単語辞書と音声情報データを切り替えることでどの話者に対しても、その話者特有の話し方をする合成音声を生成することができるようになる。図5は図4の複数の話者別の話者別単語辞書と音声情報データの組の切り替え方法の一例を示すブロック図である。テキスト解析部2、韻律情報生成部4、合成音声部5と話者別単語辞書、音声情報データの間にスイッチ10（ここでは単純化のため1個のみ示す）を設け、このスイッチ10を話者情報によって切り替える。

【0027】なお、第2の実施形態の読み変換部9を話者毎に備えてもよく、また第3の実施形態の話者別単語辞書3と読み変換部9を話者毎に備えてもよい。

【0028】以上、合成に使用する話者の発音情報を話者が実際に発した音声から抽出し、話者別単語辞書や読み変換規則を生成し、その話者の合成音声を作成する際に使用することで、元の話者の発声をより忠実に再現することが可能になる。さらに、これら話者別に話者別単語辞書や読み変換規則と音声情報データを有し、選択された話者に台わせてそれらを切り替えることで、どの合成音声に対しても元の話者の特徴を反映させることができ、合成音声のバリエーションも増やすことができるようになる。

【0029】図6は本発明の第1の実施形態をコンピュータを用いて実施した場合のブロック図である。

【0030】入力装置11はキーボード等で、テキストが入力される。記憶装置12、13はそれぞれ図1中の話者別単語辞書3、音声情報データ6に相当する。記憶装置14はハードディスクである。出力装置15は生成された合成音声出力される、スピーカ等である。記録媒体16は、図1中のテキスト解析部2、韻律情報生成部4、合成音声生成部5の各処理からなる音声合成プログラムが記録されている。FD（フロッピー・ディスク）、CD-ROM、MD（光磁気ディスク）等の記録媒体である。データ処理装置17は記録媒体16から音声合成プログラムを記憶装置14に読み込んで、これを実行するCPUである。

【0031】なお、その他の実施形態も同様にしてコンピュータを用いて実施することができる。

【0032】

【発明の効果】以上説明したように、本発明は、話者の発音に特化した辞書や変換手段を持ち、合成音声の発音

(8)

特開2000-187495

13

14

情報を元の話音と同じものにより、合成音声の個人性を向上させることが可能になる。

【図面の簡単な説明】

【図1】本発明の第1の実施の形態の音声合成装置のブロック図である。

【図2】本発明の第2の実施の形態の音声合成装置のブロック図である。

【図3】本発明の第3の実施の形態の音声合成装置のブロック図である。

【図4】本発明の第4の実施の形態の音声合成装置のブロック図である。

【図5】第4の実施形態における話音別単語辞書と音声情報データのセットの切り替え方法の一例を示す図である。

【図6】第1の実施の形態をコンピュータを用いて実行する装置のブロック図である。

【図7】音声合成装置の従来例のブロック図である。

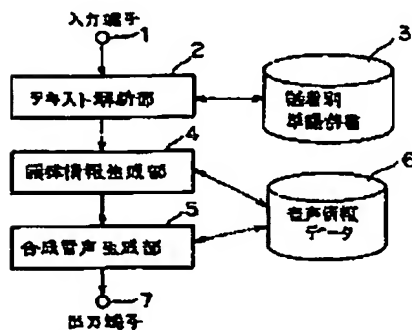
【符号の説明】

1 入力端子

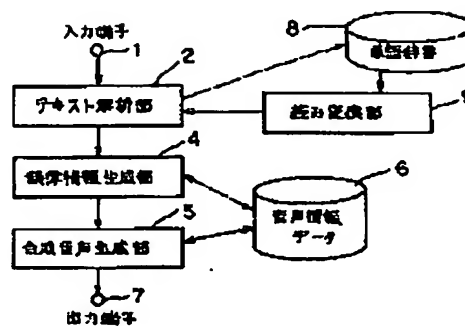
*

- * 2 テキスト解析部
- 3 3A~3C 話音別単語辞書
- 4 辞書情報生成部
- 5 合成音声生成部
- 6 6A~6C 音声情報データ
- 7 出力端子
- 8 単語辞書
- 9 読み変換部
- 10 スイッチ
- 11 入力装置
- 12, 13, 14 記憶装置
- 15 出力装置
- 16 記録媒体
- 17 データ処理装置
- 21 音声情報検索部
- 22 合成音声生成部
- 23 音声情報データベース
- 24 合成音声生成規則

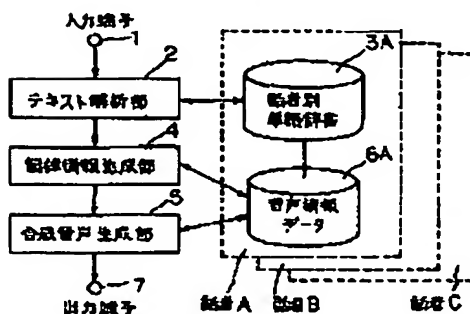
【図1】



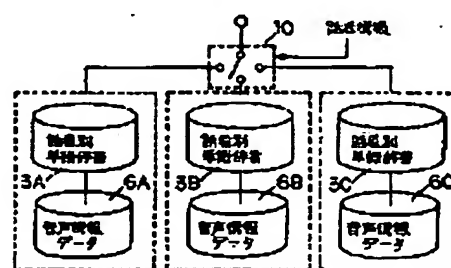
【図2】



【図4】



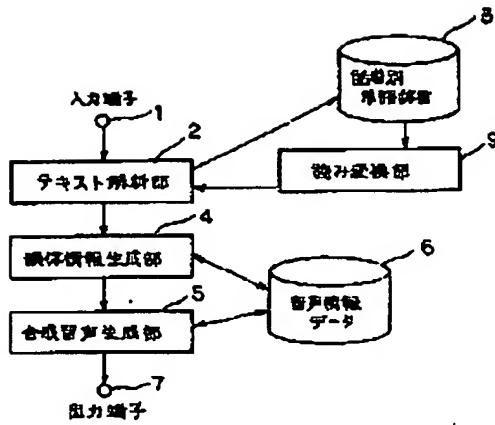
【図5】



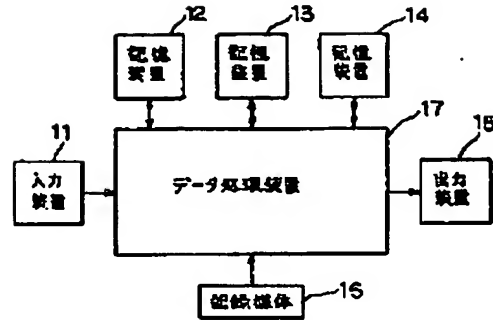
(9)

特開2000-187495

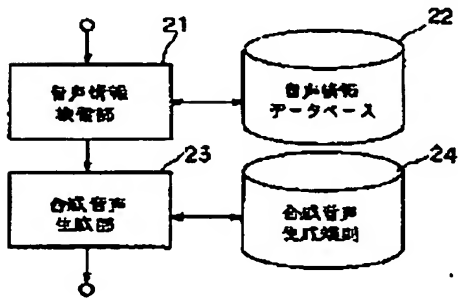
【図3】



【図6】



【図7】



**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.